Bioastronomy 2007: Molecules, Microbes, and Extraterrestrial Life
ASP Conference Series, Vol. 420, 2009
K. J. Meech, J. V. Keane, M. J. Mumma, J. L. Siefert, and D. J. Werthimer, eds.

SETI with Help from Five Million Volunteers: The Berkeley SETI Efforts

E. J. Korpela,[1] D. P. Anderson,[1] R. Bankay,[1] J. Cobb,[1] G. Foster,[1] A. Howard,[1] M. Lebofsky,[1] G. Marcy,[1] A. Parsons,[1] A. Siemion,[1] J. Von Korff,[1] D. Werthimer,[1] and K. A. Douglas[2]

[1]*University of California, Berkeley, CA, USA, 94720-7450*

[2]*School of Physics, University of Exeter, Exeter, United Kingdom*

Abstract. We summarize radio and optical SETI programs based at the University of California, Berkeley.

The ongoing SERENDIP V sky survey searches for radio signals at the 300 meter Arecibo Observatory. The currently installed configuration supports 128 million channels over a 200 MHz bandwidth with 1.6 Hz spectral resolution. Frequency stepping allows the spectrometer to cover the full 300MHz band of the Arecibo L-band receivers. The final configuration will allow data from all 14 receivers in the Arecibo L-band Focal Array to be monitored simultaneously with over 1.8 billion simultaneous channels.

SETI@home uses desktop computers volunteers to analyze over 100 TB of at taken at Arecibo. Over 5 million volunteers have run SETI@home during its 10 year history. The SETI@home sky survey is 10 times more sensitive than SERENDIP V but it covers only a 2.5 MHz band, centered on 1420 MHz. SETI@home searches a much wider parameter space, including 14 octaves of signal bandwidth and 15 octaves of pulse period with Doppler drift corrections from -100 Hz/s to +100 Hz/s.

The ASTROPULSE project is the first SETI search for μs time scale pulses in the radio spectrum. Because short pulses are dispersed by the interstellar medium, and amount of dispersion is unknown, ASTROPULSE must search through 30,000 possible dispersions. Substantial computing power is required to conduct this search, so the project will use volunteers and their personal computers to carry out the computation (using distributed computing similar to SETI@home).

The SEVENDIP optical pulse search looks for ns time scale pulses at visible wavelengths. It utilizes an automated 30 inch telescope, three ultra fast photo multiplier tubes and a coincidence detector. The target list includes F,G,K and M stars, globular cluster and galaxies.

Introduction

At the University of California, Berkeley, we are conducting five SETI searches that are roughly orthogonal to each other in search space. These five searches are summarized in Table 1.

The SERENDIP V sky survey covers a relatively broad range of radio frequencies, but not as sensitively as SETI@home. The SETI@home sky survey is more sensitive and examines a much wider variety of signal types than SERENDIP, but only covers a narrow band centered on the 21 cm Hydrogen

Program Name	Timescale	Wavelength
SERENDIP	1 s	radio
SETI@home	1 ms – 10 s	radio
ASTROPULSE	μs	radio
SEVENDIP	ns	optical

Table 1. SETI programs at the University of California, Berkeley

line (a "magic frequency"). The ASTROPULSE program is the first search for μs time scale radio pulses. The SEVENDIP optical pulse search is sensitive to low duty cycle ultra-short pulses (eg: pulsed lasers). The optical continuous search is sensitive to narrow band long duty cycle signals (eg: continuous visible lasers).

We describe each of these programs below.

Optical SETI

There is no clear wavelength choice for SETI. Microwave, IR and visible wavelengths all have advantages and disadvantages, depending on what factors another civilization chooses to optimize (power, size, bandwidth, and/or beam size). Although optical photons require more energy to generate than radio photons, optical beam sizes are typically much smaller, and directed interstellar communication links can be more efficient (Lampton 2000; Townes 1961, 1997).

SEVENDIP (Search for Extraterrestrial Visible Emissions from Nearby Developed Populations)

The SEVENDIP program at Berkeley searches for nanosecond time scale pulses, perhaps transmitted by a powerful pulsed laser operated by a distant civilization. The target list includes mostly nearby F,G,K and M stars, plus a few globular clusters and galaxies. The pulse search utilizes Berkeley's 0.8 meter automated telescope at Leuschner observatory and specialized instrumentation to detect short pulses. A similar instrument has been developed at Harvard University (Howard et al. 1999).

The SEVENDIP instrument uses beam splitters to feed light from the telescope onto three high speed photomultiplier tubes (Wright et al. 2001). These tubes have a rise time of 0.7 ns and are sensitive to 300 - 700 nm wavelengths. The three signals are fed to high speed amplifiers, fast discriminators, and a coincidence detector. Three detectors are needed to reject "false alarms," which can be caused by radioactive decay and scintillation in the PMT glass, cosmic rays, and ion feedback. These false alarms can happen often in a single PMT, but almost never occur in three PMT's simultaneously.

The Leuschner pulse search has examined several thousand stars so far, each star for one minute or more. The experiment's sensitivity is 1.5 $\times 10^{-17}$ W/m^2 for a 1 ns pulse, which corresponds to 1.5 $\times 10^{-28}$ W/m^2 average power if the pulse duty cycle is one nanosecond every 100 seconds.

The SERENDIP V Arecibo Sky Survey

The SERENDIP SETI program began 25 years ago; it has gone through four generations of instrumentation and has observed on 14 radio telescopes. During these twenty five years, SERENDIP's sensitivity has improved by a factor of ten thousand and the number of channels has increased from one hundred to more than one hundred million (Werthimer et al. 1997; Bowyer et al. 1997).

The latest SERENDIP sky survey, SERENDIP V, began in earnest in 2009. Observations are ongoing. The survey utilizes the National Astronomy and Ionospheric Center's 305 meter radio telescope in Arecibo, Puerto Rico. The survey thoroughly covers 25% of the sky (declinations from +3 to +33 degrees) and has moderate coverage from -2 to +3 and +33 to +38 degrees. Each of the 10 million beams will be observed an average of three times during the five year survey. Multiple observations are needed because sources may scintillate (Cordes, Lazio & Sagan 1997) or have short duty cycles, and many of our robust detection algorithms require multiple detections.

The sky survey utilizes a real time 128 million channel FFT spectrum analyzers to search for narrow band radio signals in a 300 MHz band centered at the 21 cm Hydrogen line (1420 MHz). The currently installed system consists of one such instrument. We are working towards a final configuration consisting of 14 of these instruments which will allow simultaneous analysis of data from all of the 14 receivers of the Arecibo L-band Focal Array (ALFA) The system has a 0.6 second integration time, 1.6 Hz resolution, and a sensitivity of 10^{-24} W/m^2.

SERENDIP V conducts observations continuously whenever the ALFA array is in uses and simultaneously with ongoing astronomy programs. SERENDIP data analysis is described by Cobb et al. (2000). Information on signals whose power exceeds 16 times the mean noise power are logged along with baseline power, telescope coordinates, time and frequency. This data is transmitted to Berkeley in real time; then, radio frequency interference (RFI) rejection algorithms are applied to the data, off-line, at UC Berkeley. After the RFI is rejected, computers search for candidate signals. SERENDIP's candidate detection algorithms are sensitive to several types of signals, which, individually or combined, may trigger an event to be noted for further study. These algorithms test for beam pattern matching, linear drift rates, regularly spaced pulses, multiple frequencies (particularly those periodic in frequency), and coincidence with nearby stars, globular clusters, or extra-solar planetary systems. Every few months, the entire data base is scanned for multiple detections – "signals" that are detected again when the telescope revisits the same sky coordinates. We test how well these multiple detections fit a barycentric reference frame. We also apply another test that allows much higher frequency separation, which is necessary if transmitters are not corrected for their planet's rotation and revolution. Data are simultaneously sent to Cornell University for analysis using other techniques.

Potential candidates are scored and ranked by the probability of noise causing that particular detection. In cases where multiple detections have been made, a joint probability is assessed. These joint probabilities are used for comparing candidates against each other and generating a prioritized candidate list for re-observation.

The SETI@home Sky Survey

SETI@home data comes from the same piggyback receiver that SERENDIP uses at the Arecibo radio telescope. Whereas SERENDIP analyzes this data primarily using a special-purpose spectrum analyzer and supercomputer located at the telescope, SETI@home records the data, and then distributes the data through the internet to hundreds of thousands of personal computers. This approach provides a tremendous amount of computing power but limits the amount of data that can be handled. Hence SETI@home covers a relatively narrow frequency range (2.5 MHz) but searches for a wider range of signal types, and with improved sensitivity (Anderson et al. 2000; Sullivan et al. 1997).

SETI@home was launched on May 17, 1999. In its 10 years of operation, it has attracted over 5 million participants. Together the participants have contributed over 2×10^{27} floating point operations making SETI@home the largest computation ever performed SETI@home is also one of the largest supercomputers on our planet, currently averaging 3.5 PFLOP actual performance. Users are located in 226 countries, and about 50% of the users are from outside the U.S.

Although SETI@home has 1/80 the frequency coverage of SERENDIP V, its sensitivity is roughly ten times better. The SETI@home search also covers a much richer variety of signal bandwidths, drift rates, and time scales than SERENDIP V or any other SETI program to date.

Primary data analysis, done using distributed computing, computes power spectra and searches for "candidate" signals such as spikes, gaussians, and pulses. Secondary analysis, done on the project's own computers, rejects RFI and searches for repeated events within the database of candidate signals.

SETI@home covers a 2.5 MHz bandwidth centered at the 1420 MHz Hydrogen line from each of the 14 ALFA receivers (7 beams $\times$ 2 polarizations). The 2.5 MHz band is recorded continuously onto SATA disks tapes with one bit complex sampling. Disks are mailed to UC Berkeley for analysis.

SETI@home data disks from the Arecibo telescope are divided into small "work units" as follows: the 2.5 MHz bandwidth data is first divided into 256 sub-bands; each work unit consists of 107 seconds of data from a given 9,765 Hz sub-band. Work units are then sent over the Internet to the client programs for the primary data analysis.

Because an extraterrestrial civilization's signal has unknown bandwidth and time scale, the client software searches for signals at 15 octave spaced bandwidths ranging from 0.075 Hz to 1220 Hz, and time scales from 0.8 ms to 13.4 seconds. The rest frame of the transmitter is also unknown (it may be on a planet that is rotating and revolving), so extraterrestrial signals are likely to be drifting in frequency with respect to the observatory's topocentric reference frame. Because the reference frame is unknown, the client software examines about 1200 different Doppler acceleration frames of rest (dubbed "chirp rates"), ranging from -100 Hz/sec to +100 Hz/sec.

At each chirp rate, peak searching is done by computing non-overlapping FFTs and their resulting power spectra. FFT lengths range from 8 to 131,072 in 15 octave steps. Peaks greater than 24 times the mean power are recorded and sent back to the SETI@home server for further analysis.

Besides searching for peaks in the multi-spectral-resolution data, SETI@home also searches for signals that match the telescope's Gaussian beam pattern. Gaussian beam fitting is computed at every frequency and every chirp rate at spectral resolutions ranging from 0.6 to 1220 Hz (temporal resolutions from 0.8 ms to 1.7 seconds). The beam fitting algorithm attempts to fit a Gaussian curve at each time and frequency in the multi-resolution spectral data. Gaussian fits whose power exceeds the mean noise power by a factor of 3.2 and whose reduced χ^2 of the gaussian fit is less than 1.42 are reported to the SETI@home servers. More details of the SETI@home analysis can be found in Korpela et al. (2001)

SETI@home also searches for pulsed signals using a modified Fast Folding Algorithm (Staelin 1969) and an algorithm which searches for three regularly-spaced pulses.

To determine signals of interest the data for each sky position which has recently received new potential signals is examined by our Near-Time Persistency Checker (NTPCkr). This program scores candidates based upon the probability that the assemblage of potentials signals seed could be due to random fluctuations in the noise background. This score included likelyhood of multiple detections in any reference frame, those that repeat in the barycentric frame, that match the antenna beam pattern, or detections coincident with newly detected planets, nearby stars (from the Hipparcos catalog) or galaxies. We generate a ranked list of our best candidates for reobservation.

Most of the signals found by the client programs turn out to be terrestrial based radio frequency interference (RFI). We employ a substantial number of algorithms to reject the several types of RFI (Cobb et al. 2000) from the best signals. Once RFI rejection has been performed on a candidate group, it is re-scored by the NTPCkr.

The SETI@home screen saver program is available for mac, windows and many versions of unix. Participants can download the client software at: `http://setiathome.berkeley.edu`.

SETI@home has clearly shown the viability of volunteer based distributed computing for other scientific problems. To this end we have developed an infrastructure dubbed BOINC (Berkeley Open Infrastructure for Network Computing). which can be used for other applications. See K. Douglas et al. this volume.

ASTROPULSE

Radio SETI searches to date have concentrated on narrowband signals as opposed to a wideband signals such as a pulses. The ASTROPULSE project is the first SETI search for μs radio pulses. ASTROPULSE detects pulse widths ranging from 1 μs to 1 ms. Such pulses might come from extraterrestrial civilizations, evaporating black holes, gamma ray bursters, certain supernovae, or pulsars. The ASTROPULSE program mines the SETI@home data archive for serendipitous detections of such events.

One of the unique features of this search is that it is the first pulse search to use coherent de-dispersion in a "blind" fashion - we have no previous knowledge of a specific dispersion measure (DM) to examine. The reason this search has never been attempted before is due to the enormous computing power required.

The computing problem is eminently parallel in nature. Similar to SETI@home, ASTROPULSE uses volunteers and their personal computers to carry out the computation. ASTROPULSE uses a general purpose distributed computing system we have developed (BOINC).

Thus far, ASTROPULSE has looked at about a year of SETI@home data, resulting in 82.2 million potential detections. We are currently working on RFI rejection techniques to clean the dataset of RFI, primarily due to aviation radars. More details of ASTROPULSE and other searches for pulsed signals that we are performing can be found in Von Korff et al. (this volume)

Acknowledgments. This work has been supported by the Planetary Society, the Seti Institute, the University of California, Sun Microsystems and donations from individuals around the planet. Key hardware was donated by Network Appliance, Xilinx, Fujifilm Computers, Toshiba, Quantum, Hewlett Packard, and Intel Corp. We receive excellent support from the staff of the Arecibo Observatory, a facility of the NSF, managed by Cornell University. SETI@home and Astropulse have been supported in part by NSF Grants AST-0808175 and AST-0307956, and NASA Grant NNX09AN69

References

Anderson, D., Werthimer, D., Cobb, J., Korpela, E., and Lebofsky, M. 2000. "SETI@home: Internet Distributed Computing for SETI," in *Bioastronomy 99: A New Era in Bioastronomy*, ed. G. A. Lemarchand and K. J. Meech, p. 511

Bowyer, S., Werthimer, D., Donnelly, C., Cobb,J., Ng, D., and Lampton, M. 1997. "Twenty Years of SERENDIP, the Berkeley SETI Effort: Past Resultsand Future Plans," in *Astronomical and Biochemical Origins and the Search for Life in the Universe*, ed. C. B. Cosmovici, S. Bowyer, and D. Werthimer, p. 667

Cobb, J., Lebofsky, M., Werthimer, D., and Bowyer, S. 2000. "SERENDIP IV: Data Acquisition, Reduction and Analysis," in *Bioastronomy 99: A New Era in Bioastronomy*, ed. G. A. Lemarchand and K. J. Meech, p. 485

Cordes, J., Lazio, T., and Sagan, C. 1997. "Scintillation-Induced Intermittency in SETI," *the Astrophysical Journal*, 487, 782

Howard, A., Horowitz, P., Coldwell, C., Latham, D., Papaliolios, C., Stefanik, R., Wolff, J., and Zajac,J. 2000. "Optical SETI at Harvard-Smithsonian" in *Bioastronomy 99: A New Era in Bioastronomy*, p. 545

Korpela, E., Werthimer, D., Anderson, D., Cobb, J. & Lebofsky, M. 2001. "SETI@home–Massively distributed computing for SETI," *Computing in Science and Engineering*, **v3n1**, 78

Lampton, M. 2000. "Optical SETI: The Next Search Frontier," in *Bioastronomy 99: A New Era in Bioastronomy*, ed. G. A. Lemarchand and K. J. Meech, p. 565

Montebugnoli, S., Monari, J., Orfei A, et al. 2000. "Setitalia - SETI in Italy," in *Bioastronomy 99: A New Era in Bioastronomy*, ed. G. A. Lemarchand and K. J. Meech, p. 501

Reines, A., and Marcy, G. 2002. "Optical SETI: A Spectroscopic Search for Laser Emission from Nearby Stars" *Publications of the Astronomical Society of the Pacific* April

Staelin, D. H. 1969. in *Proc. IEEE*, 57, 724

Stootman, F., DeHorta, A., Wellington, K., and Oliver, C. 2000. "The Southern Serendip Project," in *Bioastronomy 99: A New Era in Bioastronomy*, ed. G. A. Lemarchand and K. J. Meech, p. 491

Sullivan, W., Werthimer, D., Bowyer, S., Cobb, J., Gedye, D., and Anderson, D. 1997. "A New Major SETI Project Based on Project SERENDIP Data and 50,000 Personal Computers," in *Astronomical and Biochemical Origins and the Search for Life in the Universe*, ed. C. B. Cosmovici, S. Bowyer, and D. Werthimer, p. 729

Townes, C. 1961. "Interstellar and Interplanetary Communication by Optical Masers," in *Nature* 190, 205

Townes, C. 1997 "Optical and Infrared SETI," in *Astronomical and Biochemical Origins and the Search for Life in the Universe*, ed. C. B. Cosmovici, S. Bowyer, and D. Werthimer, p. 585

Werthimer, D., Bowyer, S., Ng, D., Donnelly, C., Cobb, J., Lampton, M., and Airieau, S. 1997. "The Berkeley SETI Program: SERENDIP IV Instrumentation," in *Astronomical and Biochemical Origins and the Search for Life in the Universe*, ed. C. B. Cosmovici, S. Bowyer, and D. Werthimer, p. 683

Wright, S., Drake, F., Stone, R., Treffers, R., and Werthimer, D. 2001 "An Improved Optical SETI Detector," SPIE proceedings on Optical SETI, ed. S. Kingsley.